

# How Search Engines Work

**S**earch engines are the most common tools used for locating information on the World Wide Web. Search engines consist of three basic but essential elements: an automatic site searcher, the index of data, and the search software.

Engines act as a sort of electronic card catalog for Web pages on the Internet. They collect and store data about millions of Web pages and then offer users a handy interface with which they can search for information based on keywords.

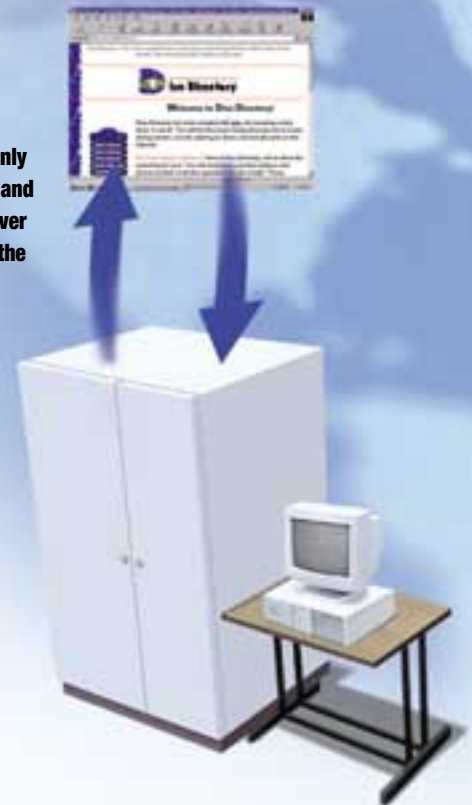
While each search engine collects and stores Web site data in its own manner, the most common method for collecting information is through the use of bots, also known as robots, spiders, and crawlers. (Please see page 194 for more information on spiders and crawlers.) These bots are the automatic site searchers that are sent out by servers to crawl through millions of Web sites and index all or some of the words on each Web page.

Each search engine's bots performs their crawls differently. Some bots search daily, some weekly, and some monthly. They also vary in the degree of crawl they perform. This is illustrated to the right.

The data collected by the bots is then indexed on the server so users can access the information through keyword searches. In this illustration, you can see how the search software reads the words on the page to locate matches to the keyword search performed.

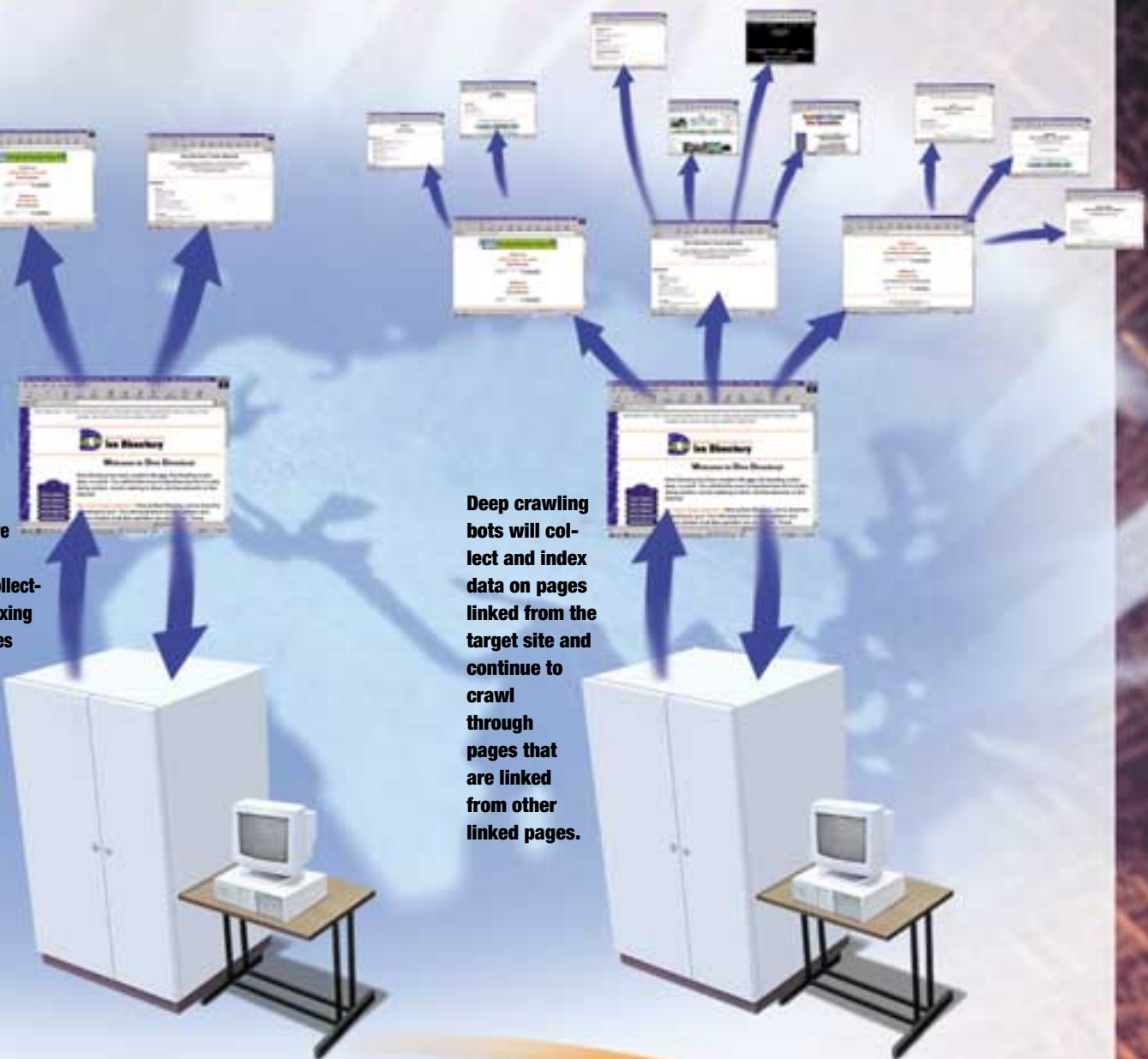
*Compiled by Lori Robison  
Graphics & Design by Lori Garriss*

**Some bots go out to only one specific Web site and return back to the server immediately to index the site contents.**



Other bots perform more thorough searches, collecting and indexing data on pages linked from the target site.

Deep crawling bots will collect and index data on pages linked from the target site and continue to crawl through pages that are linked from other linked pages.



When a user performs a keyword search in a search engine, the engine searches its index for pages that include the requested keywords. The search software counts the number of times the keyword or keywords appear on the page and in the page's Meta tags to determine the page's relevancy.



```
<html>
<head>
<title>Dive Directory - World Wide Scuba Diving Directory - Dive
Centers - Dive Resorts -
Dive Charters </title>
<meta NAME="description" CONTENT="Dive Directory - The most
comprehensive scuba diving directory online! Complete index for dive
centers, dive resorts and dive charters world wide!">
<meta NAME="keywords" CONTENT="dive directory, scuba diving
directory,
diving directory, world wide scuba diving directory">
<meta NAME="Author" CONTENT="info@divedirectory.com
&lt;mailto:CONTENT-&gt;">
<meta NAME="Distribution" Content="Global">
```